

On vector spaces of linearizations for matrix polynomials in orthogonal bases

Heike Faßbender^a, Philip Saltenberger^{*,a}

^a*Institut Computational Mathematics/ AG Numerik, TU Braunschweig, Pockelsstr. 14,
38106 Braunschweig, Germany*

Abstract

Following the ideas in [8], for matrix polynomials $P(\lambda) = \sum_{i=0}^k P_i \phi_i(\lambda)$, $P_i \in \mathbb{R}^{n \times n}$ given in an orthogonal basis $\phi_0, \phi_1, \dots, \phi_k$, the corresponding vector spaces, called $\mathbb{M}_1(P)$, $\mathbb{M}_2(P)$ and $\mathbb{DM}(P)$, of potential linearizations for $P(\lambda)$ are analyzed. All pencils in $\mathbb{M}_1(P)$ are characterized concisely. Moreover, an easy to check criterion whether a pencil in $\mathbb{M}_1(P)$ is a (strong) linearization of $P(\lambda)$ is given. Results on the vector space dimensions, the genericity of linearizations in $\mathbb{M}_1(P)$ and the form of block-symmetric pencils are derived in a new way on a basic algebraic level. Throughout the paper, structural resemblances between the matrix pencils in \mathbb{L}_1 , i.e. the results obtained in [8], and their generalized versions are pointed out.

Key words: matrix polynomial, (strong) linearization, orthogonal basis, block-symmetry, ansatz space, structure-preserving linearization
2010 MSC: 65F15, 15A03, 15A18, 15A22

1. Introduction

Linearization of matrix polynomials expressed in standard and nonstandard bases have received much attention in recent years. In the groundbreaking paper [8] vector spaces of possible linearizations of matrix polynomials have been introduced. These turned out to build an elegant framework to find and construct linearizations for square matrix polynomials as well as to study their algebraical and analytical properties. While the paper [8] is mainly concerned with the characterization and analysis of these spaces for

^{*}Corresponding author, Email philip.saltenberger@tu-braunschweig.de

matrix polynomials in the standard monomial basis, recently the research on matrix polynomials and linearizations expressed in nonstandard polynomial bases has received more attention, see, e.g., [1, 3, 5, 4, 7, 9, 10, 11].

This paper is devoted to the study of (regular) matrix polynomials $P(\lambda) = \sum_{i=0}^k P_i \phi_i(\lambda)$, $P_i \in \mathbb{R}^{n \times n}$ expressed in an orthogonal basis $\{\phi_i(\lambda)\}_{i=0}^k$, generalizing most concepts from [8] to this special case. In particular, we will consider the set $\mathbb{M}_1(P)$ of all $kn \times kn$ matrix pencils $\mathcal{L}(\lambda)$ satisfying

$$\mathcal{L}(\lambda)(\Phi_k(\lambda) \otimes I_n) = v \otimes P(\lambda)$$

with $\Phi_k(\lambda) := [\phi_{k-1} \dots \phi_1 \ \phi_0]^T$. For the standard basis, this is just the definition of $\mathbb{L}_1(P)$ [8, Definition 3.1] with $\Phi_k(\lambda) = [\lambda^{k-1} \dots \lambda \ 1]^T =: \Lambda_k$. The same kind of generalization of $\mathbb{L}_1(P)$ to matrix polynomials in nonstandard bases has been already considered, e.g., in [9, 3]. We will give an explicit characterization of the elements of $\mathbb{M}_1(P)$ that enables us to formulate our results readily accessible providing quite short proofs. Moreover, we show how to easily construct linearizations by means of an intuitive and readily checked linearization condition. Clearly, most of our findings are equivalent to already known results. Thus our main contribution here is a new view aiming to open up new perspectives on the structure of ansatz spaces in general and present even well-known facts in a new livery. A second main goal is to present the facts in a concise and succinct manner keeping the proofs on a basic algebraic level without drawing on deeper theoretical results.

In Section 2 the basic notation used and some well-known results are summarized. In Section 3, generalized ansatz spaces for orthogonal bases are defined and their basic properties are proven. The extension of the double ansatz space from [8] to orthogonal bases is the subject of Section 4 whereas Section 5 provides a construction algorithm for block-symmetric pencils. Section 6 presents some concluding remarks.

2. Preliminaries and Basic Notation

For $\mathbb{R}[\lambda]$, the ring of real polynomials in the variable λ , the $n \times n$ matrix ring over $\mathbb{R}[\lambda]$ is denoted by $\mathbb{R}[\lambda]^{n \times n}$. Its elements are referred to as matrix polynomials. Notice that $\mathbb{R}[\lambda]^{n \times n}$ is a vector space over \mathbb{R} . We consider matrix polynomials $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ expressed in polynomial bases $\Phi = \{\phi_j(\lambda)\}_{j=0}^\infty$ that follow a three-term recurrence relation. In particular we assume that

$$\alpha_j \phi_{j+1}(\lambda) = (\lambda - \beta_j) \phi_j(\lambda) - \gamma_j \phi_{j-1}(\lambda) \quad j \geq 1 \quad (1)$$

for some coefficients $\alpha_j \neq 0, \beta_j, \gamma_j \in \mathbb{R}$ and $\phi_{-1}(\lambda) = 0, \phi_0(\lambda) = 1$. Popular special cases include the monomials, Newton and Chebyshev bases or the Legendre basis. Moreover, we usually assume that $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ may be expressed as

$$P(\lambda) = P_k \phi_k(\lambda) + P_{k-1} \phi_{k-1}(\lambda) + \dots + P_1 \phi_1(\lambda) + P_0 \phi_0(\lambda) \quad (2)$$

with $P_k \neq 0$. In this case $P(\lambda)$ is said to have degree k , i.e. $\deg(P(\lambda)) = k$. A matrix polynomial with $\det(P(\lambda)) \neq 0$ is called regular, otherwise it is called singular. Moreover, matrix polynomials of degree one are called matrix pencils.

Any scalar $\alpha \in \mathbb{C}$ such that $P(\alpha) \in \mathbb{C}^{n \times n}$ is singular is called a finite eigenvalue of $P(\lambda)$. The corresponding eigenspace is defined to be $\text{null}(P(\alpha))$, i.e. the nullspace of $P(\alpha)$. Whenever $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ has degree k , the reversal of $P(\lambda)$ is the matrix polynomial

$$\text{rev}_k(P(\lambda)) := \lambda^k P\left(\frac{1}{\lambda}\right)$$

of which it can be proven that its nonzero finite eigenvalues are the reciprocals of those of $P(\lambda)$. Moreover, if zero is an eigenvalue of $\text{rev}_k(P(\lambda))$, we say that ∞ is an eigenvalue of $P(\lambda)$.

Assume $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ has degree k . Then a $kn \times kn$ matrix pencil $\mathcal{L}(\lambda) = X\lambda + Y$ is called a linearization for $P(\lambda)$ if there exist two matrix polynomials $U(\lambda), V(\lambda) \in \mathbb{R}[\lambda]^{kn \times kn}$ with nonzero, real determinants such that

$$U(\lambda)\mathcal{L}(\lambda)V(\lambda) = \left[\begin{array}{c|ccc} P(\lambda) & 0_n & \cdots & 0_n \\ \hline 0_n & & & \\ \vdots & & I_{(k-1)n} & \\ 0_n & & & \end{array} \right].$$

holds. Here I_n denotes the $n \times n$ identity matrix whereas 0_n is the $n \times n$ matrix of all zeros. A linearization $\mathcal{L}(\lambda)$ for $P(\lambda)$ is called strong whenever $\text{rev}_1(\mathcal{L}(\lambda))$ is a linearization for $\text{rev}_k(P(\lambda))$ as well. In case $\mathcal{L}(\lambda)$ is a strong linearization of a matrix polynomial $P(\lambda)$, $\mathcal{L}(\lambda)$ and $P(\lambda)$ share the same eigenvalues with the same algebraic and geometric multiplicities. Moreover, if V is a regular square matrix of appropriate dimension and $\mathcal{L}(\lambda)$ is a strong linearization, then $V\mathcal{L}(\lambda)$ is a strong linearization as well. The matrix pencils $V\mathcal{L}(\lambda)$ and $\mathcal{L}(\lambda)$ are usually called (strongly) equivalent.

Whenever a $kn \times kn$ matrix pencil $\mathcal{L}(\lambda)$ may be expressed as

$$\mathcal{L}(\lambda) = \sum_{i,j=1}^k e_i e_j^T \otimes \mathcal{L}_{ij}(\lambda) \quad (3)$$

for certain $n \times n$ matrices $\mathcal{L}_{ij}(\lambda)$, we call $\mathcal{L}(\lambda)^{\mathcal{B}} = \sum_{i,j=1}^k e_j e_i^T \otimes \mathcal{L}_{ij}(\lambda)$ the block-transpose of $\mathcal{L}(\lambda)$ (see [6, Definition 2.1]). Therefore, if $\mathcal{L}(\lambda)$ of the form (3) satisfies $\mathcal{L}(\lambda) = \mathcal{L}(\lambda)^{\mathcal{B}}$ it is called block-symmetric, whereas it is called block-skew-symmetric whenever $\mathcal{L}(\lambda) = -\mathcal{L}(\lambda)^{\mathcal{B}}$.¹ For the $s \times s$ leading principal submatrix of a matrix polynomial $P(\lambda)$ we use the notation $[P(\lambda)]_s$. Using MATLAB notation this means $[P(\lambda)]_s = (P(\lambda))(1:s, 1:s)$.

3. Generalized Ansatz Spaces

Whenever this is not further specified, in this section $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ is without exception assumed to be a regular matrix polynomial expressed in an orthogonal basis as in (2) with $\deg(P(\lambda)) = k \geq 2$.² The assumption of regularity is made since our results rely in large parts on [1, Sec. 2.1] where only regular matrix polynomials have been considered.

For $P(\lambda)$ as in (2) we define $\Phi_k(\lambda) := [\phi_{k-1}(\lambda) \ \dots \ \phi_1(\lambda) \ \phi_0(\lambda)]^T$ and consider the set $\mathbb{M}_1(P)$ of all $kn \times kn$ matrix pencils $\mathcal{L}(\lambda)$ satisfying

$$\mathcal{L}(\lambda)(\Phi_k(\lambda) \otimes I_n) = v \otimes P(\lambda) \quad (4)$$

for some “ansatz vector” $v \in \mathbb{R}^k$. For the standard basis, this is just the definition of $\mathbb{L}_1(P)$ [8, Definition 3.1] with $\Phi_k(\lambda) = [\lambda^{k-1} \ \dots \ \lambda \ 1]^T =: \Lambda_k$. The same kind of generalization of $\mathbb{L}_1(P)$ to matrix polynomials in nonstandard bases has been considered, e.g., in [9, 3].

Certainly, $\mathbb{M}_1(P)$ is a \mathbb{R} -vector space. Next, we introduce the $n \times kn$ rectangular matrix pencil

$$m_P(\lambda) := \begin{bmatrix} \frac{(\lambda - \beta_{k-1})}{\alpha_{k-1}} P_k + P_{k-1} & P_{k-2} - \frac{\gamma_{k-1}}{\alpha_{k-1}} P_k & P_{k-3} & \cdots & P_1 & P_0 \end{bmatrix}.$$

¹Notice that a nonzero skew-symmetric matrix polynomial $P(\lambda)$ (that is, $P(\lambda) = -P(\lambda)^T$) of the form (3) is never block-skew-symmetric.

²We make this assumption to avoid the potential occurrence of pathological cases. Furthermore, the main purpose of this paper is to construct linearizations for $P(\lambda)$ which is superfluous when $P(\lambda)$ is already linear.

It is easily seen that $m_P(\lambda)(\Phi_k(\lambda) \otimes I_n) = P(\lambda)$. Moreover, for the $k-1 \times k$ matrix pencil

$$M_P^*(\lambda) = \begin{bmatrix} -\alpha_{k-2} & (\lambda - \beta_{k-2}) & -\gamma_{k-2} & & & \\ & -\alpha_{k-3} & (\lambda - \beta_{k-3}) & -\gamma_{k-3} & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\alpha_1 & (\lambda - \beta_1) & -\gamma_1 \\ & & & & -\alpha_0 & (\lambda - \beta_0) \end{bmatrix}$$

we have $M_P^*(\lambda)\Phi_k(\lambda) = 0$. Now we define

$$M_P(\lambda) := M_P^*(\lambda) \otimes I_n.$$

Certainly $M_P(\lambda)(\Phi_k(\lambda) \otimes I_n) = 0$ holds. We set

$$F_\Phi^P(\lambda) := \begin{bmatrix} m_p(\lambda) \\ M_P(\lambda) \end{bmatrix} \in \mathbb{R}[\lambda]^{kn \times kn}.$$

By construction

$$F_\Phi^P(\lambda)(\Phi_k(\lambda) \otimes I_n) = e_1 \otimes P(\lambda),$$

thus, $F_\Phi^P(\lambda) \in \mathbb{M}_1(P)$ with ansatz vector $e_1 \in \mathbb{R}^k$. According to [1, Thm. 2] $F_\Phi^P(\lambda)$ is a strong linearization for any regular $P(\lambda)$. In fact, $F_\Phi^P(\lambda)$ may be utilized as an “anchor pencil” to construct $\mathbb{M}_1(P)$. To this end, the next theorem gives a concise and succinct characterization of $\mathbb{M}_1(P)$ for any regular matrix polynomial $P(\lambda)$ expressed in some orthogonal polynomial basis.

Theorem 1 (Characterization of $\mathbb{M}_1(P)$). *Let $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ be regular and of degree $\deg(P(\lambda)) = k \geq 2$. Then $\mathcal{L}(\lambda) \in \mathbb{M}_1(P)$ if and only if*

$$\mathcal{L}(\lambda) = [(v \otimes I_n) \ B] F_\Phi^P(\lambda) \tag{5}$$

for some vector $v \in \mathbb{R}^k$ and some matrix $B \in \mathbb{R}^{kn \times (k-1)n}$.

Proof. It is immediate that any matrix pencil $\mathcal{L}(\lambda) = [(v \otimes I_n) \ B] F_\Phi^P(\lambda)$ satisfies (4) since

$$\begin{aligned} [(v \otimes I_n) \ B] F_\Phi^P(\lambda) (\Phi_k(\lambda) \otimes I_n) &= [(v \otimes I_n) \ B] (e_1 \otimes P(\lambda)) \\ &= v \otimes P(\lambda). \end{aligned}$$

Now let $\mathcal{L}(\lambda) \in \mathbb{M}_1(P)$, thus, $\mathcal{L}(\lambda)(\Phi_k(\lambda) \otimes I_n) = v \otimes P(\lambda)$ has to hold. As $v \otimes P(\lambda) = \sum_{i=0}^k (v \otimes P_i \phi_i(\lambda))$ it follows that $\mathcal{L}(\lambda)(\Phi_k(\lambda) \otimes I_n)$ has to generate the term $v \otimes P_k \phi_k(\lambda)$ on the right hand side of (4). Further, the recurrence relation (1)

$$v \otimes P_k \phi_k(\lambda) = v \otimes (\alpha_{k-1}^{-1}((\lambda - \beta_{k-1})\phi_{k-1}(\lambda) - \gamma_{k-1}\phi_{k-2}(\lambda))P_k)$$

gives that $\mathcal{L}(\lambda)$ may be expressed as

$$\mathcal{L}(\lambda) = [(v \otimes \alpha_{k-1}^{-1}P_k) \mathcal{L}_1] \lambda + [l^* \mathcal{L}_0]$$

for some matrices $l^* \in \mathbb{R}^{kn \times n}$ and $\mathcal{L}_1, \mathcal{L}_0 \in \mathbb{R}^{kn \times (k-1)n}$. Now observe that $\mathcal{L}^*(\lambda) := [(v \otimes I_n) \mathcal{L}_1] F_\Phi^P(\lambda)$ has the form

$$\mathcal{L}^*(\lambda) = [(v \otimes \alpha_{k-1}^{-1}P_k) \mathcal{L}_1] \lambda + [(v \otimes I_n) \mathcal{L}_1] F_\Phi^P(0)$$

as

$$F_\Phi^P(\lambda) = F_\Phi^P(0) + \begin{bmatrix} \frac{\lambda}{\alpha_{k-1}} P_k & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \lambda I_{k-1} & \\ 0 & & & \end{bmatrix}.$$

Thus $\Delta \mathcal{L}(\lambda) := \mathcal{L}(\lambda) - \mathcal{L}^*(\lambda) \in \mathbb{R}^{kn \times kn}$, i.e. it is independent of λ . Moreover, $\Delta \mathcal{L}(\lambda)$ satisfies $\Delta \mathcal{L}(\lambda)(\Phi_k(\lambda) \otimes I_n) = 0$. Since $\phi_0(\lambda), \dots, \phi_{k-1}(\lambda), \lambda \phi_{k-1}(\lambda)$ form a basis of $\mathbb{R}_k[\lambda]$, the vector space of real polynomials of degree $\leq k$, this implies $\Delta \mathcal{L} = 0$ and proves that $\mathcal{L}(\lambda) = \mathcal{L}^*(\lambda)$. \square

In other words, Theorem 1 states that

$$\mathbb{M}_1(P) = \{ [(v \otimes I_n) B] F_\Phi^P(\lambda) \mid v \in \mathbb{R}^k, B \in \mathbb{R}^{kn \times (k-1)n} \}.$$

In case, $\Phi_k = \Lambda_k$ denotes the monomial basis, $F_\Phi^P(\lambda)$ is just the first Frobenius companion form for $P(\lambda)$ [8, (3.1)] and $\mathbb{M}_1(P) = \mathbb{L}_1(P)$. The description of $\mathbb{L}_1(P)$ in [8, Lem. 3.4, Thm. 3.5] differs from (5) significantly although both characterizations are easily seen to be equivalent. Furthermore, since any pencil $\mathcal{L}(\lambda) = [(v \otimes I_n) B] F_\Phi^P(\lambda)$ can be uniquely identified with the tuple (v, B) we obtain the isomorphism

$$\mathbb{M}_1(P) \cong \mathbb{R}^k \times \mathbb{R}^{kn \times (k-1)n}.$$

This isomorphism was also observed in the proof of [3, Thm. 4.4] in the context of matrix polynomials in standard basis.

Corollary 1. *For any regular $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ of degree k*

$$\dim(\mathbb{M}_1(P)) = k(k-1)n^2 + k.$$

Corollary 1 is essentially [8, Cor. 3.6] for the monomial basis; $\Phi_k(\lambda) = \Lambda_k$. We now give a universal linearization condition for matrix pencils in $\mathbb{M}_1(P)$ that does not depend on the chosen basis at all.

Corollary 2. *Let $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ be regular. Then $\mathcal{L}(\lambda) \in \mathbb{M}_1(P)$ in (5) is a strong linearization for $P(\lambda)$ if and only if*

$$\text{rank}([(v \otimes I_n) \ B]) = kn. \quad (6)$$

Certainly, (6) is equivalent to $[(v \otimes I_n) \ B] \in \text{GL}_{kn}(\mathbb{R})$.

Proof. Whenever $\text{rank}([(v \otimes I_n) \ B]) = kn$, $\mathcal{L}(\lambda) = [(v \otimes I_n) \ B]F_\Phi^P(\lambda)$ is strongly equivalent to $F_\Phi^P(\lambda)$ and thus a strong linearization for $P(\lambda)$. On the other hand, if $\text{rank}([(v \otimes I_n) \ B]) < kn$, $\mathcal{L}(\lambda)$ is singular and therefore not a linearization for the regular $P(\lambda)$. \square

For the monomial basis, Corollary 2 is essentially just a reformulation of [8, Thm. 4.1]. However, the condition for checking whether a linearization is strong given in [8, Thm. 4.1] requires that the linearization in $\mathbb{L}_1(P)$ is expressed with the ansatz vector e_1 . Fortunately, we may apply Corollary 2 to the pencils in $\mathbb{M}_1(P)$ (given as in (5)) right away without changing the ansatz vector. The construction of strong linearizations for matrix polynomials expressed in the Chebyshev basis proposed in [7] gets along without such conditions.

Note that in [8, Thm. 4.3] the three equivalent conditions

$$\begin{aligned} L(\lambda) \text{ is a linearization for } P(\lambda) &\Leftrightarrow L(\lambda) \text{ is a regular pencil} \Leftrightarrow \\ L(\lambda) &\text{ is a strong linearization for } P(\lambda) \end{aligned}$$

are given where $P(\lambda)$ is a regular matrix polynomial in the monomial basis and $L(\lambda) \in \mathbb{L}_1(P)$. These three equivalent conditions are derived in [9, Thm. 2.1] for regular matrix polynomials expressed in any polynomial bases. In fact, all three equivalent conditions are equivalent to (6) for regular matrix polynomials expressed in orthogonal bases. Corollary 2 may be utilized to formulate a simple proof of the Strong Linearization Theorem [9, Thm. 2.1] for orthogonal bases without the performance of a basis change. For a discussion

for linearizations of singular polynomials in non-monomial bases see [3, Sec. 7].

Since almost every matrix of the form $[(v \otimes I_n) \ B]$ has full rank, we obtain the following genericity statement. This result was already stated in [8, Thm. 4.7] for matrix polynomials $P(\lambda)$ in monomial basis.

Corollary 3. *For any regular $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ almost every matrix pencil in $\mathbb{M}_1(P)$ is a strong linearization for $P(\lambda)$.*

We now show how eigenvectors for $P(\lambda)$ as in (2) may be recovered from eigenvectors of linearizations in $\mathbb{M}_1(P)$. The main ideas behind this derivation follow exactly the approach in [8]. Assume $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ to be regular. Then, for a strong linearization $\mathcal{L}(\lambda) = [(v \otimes I_n) \ B]F_\Phi^P(\lambda)$, an eigenvalue $\alpha \in \mathbb{C}$ and an eigenvector $u \in \mathbb{C}^{kn}$ corresponding to α it is immediate that $\mathcal{L}(\alpha)u = 0$ if and only if $F_\Phi^P(\alpha)u = 0$ due to the regularity of $[(v \otimes I_n) \ B]$. Moreover, since for $w \in \mathbb{C}^n$

$$F_\Phi^P(\alpha)(\Phi_k(\alpha) \otimes w) = e_1 \otimes P(\alpha)w (= 0)$$

due to (4), $\Phi_k(\alpha) \otimes w$ is an eigenvector of $F_\Phi^P(\lambda)$ with corresponding eigenvalue α whenever w is an eigenvector of $P(\lambda)$ for the same eigenvalue.

Proposition 1. *Let $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ of the form (2) be regular and of degree $\deg(P(\lambda)) = k$. Furthermore, let $\mathcal{L}(\lambda) \in \mathbb{M}_1(P)$ be a linearization for $P(\lambda)$.*

1. *Let $\alpha \in \mathbb{C}$ be a finite eigenvalue of $P(\lambda)$. Then $u \in \mathbb{C}^{kn}$ satisfies $\mathcal{L}(\alpha)u = 0$ if and only if $u = \Phi_k(\alpha) \otimes w$ for some $w \in \text{null}(P(\alpha))$.*
2. *Let $\alpha = \infty$. Then $y \in \mathbb{C}^{kn}$ satisfies $\text{rev}_1(\mathcal{L}(0))y = 0$ if and only if $y = \beta e_1 \otimes w$ for some $w \in \text{null}(\text{rev}_k(P(0)))$ and $\beta \in \mathbb{C}$.*

Proof. 1. Let $P(\lambda)$ be regular and $\alpha \in \mathbb{C}$ be a finite eigenvalue of $P(\lambda)$. Assume that $u_1, \dots, u_d \subset \mathbb{C}^n$ form a basis of $\text{null}(P(\alpha))$. Then $\dim(\text{null}(\mathcal{L}(\alpha))) = d$. Therefore, since $\Phi_k(\alpha) \otimes u_1, \dots, \Phi_k(\alpha) \otimes u_d \subset \mathbb{C}^{kn}$ are easily seen to be linear independent they form a basis of $\text{null}(\mathcal{L}(\alpha))$. Now any linear combination of $\Phi_k(\alpha) \otimes u_1, \dots, \Phi_k(\alpha) \otimes u_d$ is again of the form $\Phi_k(\alpha) \otimes u$ for some vector $u \in \text{null}(P(\alpha))$.

2. Since $\text{rev}_k(\mathcal{L}(\lambda)(\Phi_k(\lambda) \otimes I_n)) = \text{rev}_1(\mathcal{L}(\lambda))\text{rev}_{k-1}(\Phi_k(\lambda) \otimes I_n)$, we have from (4)

$$\text{rev}_1(\mathcal{L}(\lambda))(\text{rev}_{k-1}(\Phi_k(\lambda)) \otimes I_n) = v \otimes \text{rev}_k(P(\lambda)).$$

Thus, if $\text{rev}_k(P(0))u = 0$ we obtain $\text{rev}_1(\mathcal{L}(0))(\text{rev}_{k-1}(\Phi_k(0)) \otimes u) = 0$. Now $\text{rev}_{k-1}(\Phi_k(0)) = \beta e_1 \in \mathbb{R}^k$ for some $\beta \in \mathbb{R}$. Since $\mathcal{L}(\lambda)$ is a strong linearization, we have $\dim(\text{null}(\text{rev}_k(P(0)))) = \dim(\text{null}(\text{rev}_1(\mathcal{L}(0))))$ and the same argumentation as in the finite case gives the statement. \square

The eigenvector recovery for the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ can be found in [8, Sec. 3]. In the special case of singular matrix polynomials this was considered in [3, Sec. 5] in the context of minimal indices.

4. Double Generalized Ansatz Spaces and Block-Symmetry

Beside (4) we may consider its transposed version

$$(\Phi_k(\lambda)^T \otimes I_n) \mathcal{L}(\lambda) = v^T \otimes P(\lambda) = \sum_{i=0}^k (v^T \otimes P_i \phi_i(\lambda)). \quad (7)$$

As before, all matrix pencils satisfying (7) form an \mathbb{R} -vector space, which we denote by $\mathbb{M}_2(P)$. For the monomial basis $\mathbb{M}_2(P) = \mathbb{L}_2(P)$, see [8, Def. 3.9]. It is characterized analogously to Theorem 1.

Theorem 2 (Characterization of $\mathbb{M}_2(P)$). *Let $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ be regular and of degree $\deg(P(\lambda)) = k \geq 2$. Then $\mathcal{L}(\lambda) \in \mathbb{M}_2(P)$ if and only if*

$$\mathcal{L}(\lambda) = F_{\Phi}^P(\lambda)^{\mathcal{B}} \begin{bmatrix} v^T \otimes I_n \\ B^{\mathcal{B}} \end{bmatrix} \quad (8)$$

for some vector $v \in \mathbb{R}^k$ and some matrix $B \in \mathbb{R}^{kn \times (k-1)n}$.

Certainly, if $\mathcal{L}(\lambda)$ satisfies (4), $\mathcal{L}(\lambda)^{\mathcal{B}}$ satisfies (7) and vice versa. Consequently, if $\mathcal{L}(\lambda) = \mathcal{L}(\lambda)^{\mathcal{B}}$, $\mathcal{L}(\lambda) \in \mathbb{M}_1(P) \cap \mathbb{M}_2(P)$. Thus, the \mathbb{R} -vector space $\mathbb{DM}(P) := \mathbb{M}_1(P) \cap \mathbb{M}_2(P)$ contains all block-symmetric pencils from $\mathbb{M}_1(P)$. Similarly, in the monomial case, the double ansatz space $\mathbb{DL}(P) = \mathbb{L}_1(P) \cap \mathbb{L}_2(P)$ contains all block-symmetric pencils from $\mathbb{L}_1(P)$, see [6].

Notice that $\mathbb{M}_1(P) \cong \mathbb{M}_2(P)$ which is easily seen by (5) and (8). This was shown for $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ in [8]. We now give a surprising statement on block-skew-symmetric pencils in $\mathbb{M}_1(P)$.

Proposition 2. *Let $\mathcal{L}(\lambda) \in \mathbb{M}_1(P)$ be block-skew-symmetric. Then $\mathcal{L}(\lambda)$ satisfying (4) with $v = [0 \ v_2 \ v_3 \ \dots \ v_k]^T \in \mathbb{R}^k$ implies $\mathcal{L}(\lambda) \equiv 0$.*

Proof. Let

$$\begin{aligned}\mathcal{L}(\lambda) &= [(v \otimes I_n) B] F_{\Phi}^P(\lambda) = [(v \otimes \alpha_{k-1}^{-1} P_k) B] \lambda + [(v \otimes I_n) B] F_{\Phi}^P(0) \\ &= - \begin{bmatrix} v^T \otimes \alpha_{k-1}^{-1} P_k \\ B^{\mathcal{B}} \end{bmatrix} \lambda - F_{\Phi}^P(0)^{\mathcal{B}} \begin{bmatrix} v^T \otimes I_n \\ B^{\mathcal{B}} \end{bmatrix}\end{aligned}\quad (9)$$

be block-skew-symmetric and assume $v = [0 \ v_2 \ v_3 \ \cdots \ v_k]^T$. Regarding (9), the block-skew-symmetry of $\mathcal{L}(\lambda)$ a priori implies B to have the form

$$B = \begin{bmatrix} Z \\ B^{\star} \end{bmatrix}$$

with $Z = -\alpha_{k-1}^{-1} P_k \otimes [v_2 \ v_3 \ \cdots \ v_k] \in \mathbb{R}^{n \times (k-1)n}$ and a block-skew-symmetric $(k-1)n \times (k-1)n$ matrix B^{\star} . Let $B^{\star} = [B_{i,j}^{\star}]_{i,j=1}^{k-1}$ with $B_{i,j}^{\star} \in \mathbb{R}^{n \times n}$. The block-skew-symmetry then implies $B_{j,j}^{\star} = 0_n$ for all $j = 1, \dots, k-1$. Now we consider the leading principal submatrices of $\mathcal{L}(\lambda)$ which certainly all have to be block-skew-symmetric.

Since $[\mathcal{L}(\lambda)]_n = \alpha_{k-2} \alpha_{k-1}^{-1} v_2 P_k \stackrel{!}{=} 0$, we have $v_2 = 0$. Now choose an index $2 \leq i \leq k-1$ and assume $v_1 = v_2 = \dots = v_i = 0$ and $[B^{\star}]_{(i-1)n} = 0$.³ Then the $in \times in$ leading principal submatrix $[\mathcal{L}(\lambda)]_{in}$ of $\mathcal{L}(\lambda)$ takes in absolute value $|[\mathcal{L}(\lambda)]_{in}|$ the form

$$|[\mathcal{L}(\lambda)]_{in}| = \left| \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \alpha_{k-1-i} \alpha_{k-1}^{-1} v_{i+1} P_k \\ \alpha_{k-1-i} B_{1,i} \\ \vdots \\ \alpha_{k-1-i} B_{i-1,i} \end{bmatrix} \right|.$$

Since $[\mathcal{L}(\lambda)]_{in}$ is block-skew-symmetric it follows that $B_{1,i} = \dots = B_{i-1,i} = 0_n$ and in particular $v_{i+1} = 0$. Therefore we have shown that $v_1 = \dots = v_{i+1} = 0$ and that $[B^{\star}]_{in} = 0$. Inductively, $i = k-1$ yields $v = 0$ and $B^{\star} = 0$. \square

Using Proposition 2 we obtain an easy proof of the following theorem.

Theorem 3. *Any matrix pencil $\mathcal{L}(\lambda) \in \mathbb{DM}(P)$ is block-symmetric.*

³Notice that these conditions are satisfied for $i = 2$.

Proof. Let $\mathcal{L}(\lambda) \in \mathbb{DM}(P)$. Then $\mathcal{L}(\lambda)$ can be expressed as

$$\begin{aligned}\mathcal{L}(\lambda) &= v \otimes m_P(\lambda)^T + B_1 M_P(\lambda) \\ &= w^T \otimes m_P(\lambda) + M_P(\lambda)^B B_2^B\end{aligned}$$

as an element of $\mathbb{M}_1(P)$ and $\mathbb{M}_2(P)$ respectively. Regarding $\mathcal{L}(\lambda)$ in the form $\mathcal{L}(\lambda) = X\lambda + Y$ this shows that $[X]_n = v_1 \alpha_{k-1}^{-1} P_k = w_1 \alpha_{k-1}^{-1} P_k$. Thus it follows that $v_1 = w_1$. Now note that $\mathcal{L}(\lambda)$ (seen as an element of $\mathbb{M}_2(P)$) via block-transposition becomes an element of $\mathbb{M}_1(P)$. Therefore

$$\begin{aligned}\tilde{\mathcal{L}}(\lambda) &:= \mathcal{L}(\lambda) - \mathcal{L}(\lambda)^B = (v - w) \otimes m_P(\lambda) + (B_1 - B_2) M_P(\lambda) \\ &=: \tilde{v} \otimes m_P(\lambda) + \tilde{B} M_P(\lambda)\end{aligned}$$

is a block-skew-symmetric pencil in $\mathbb{M}_1(P)$. Since $\tilde{v} = [0 \ \tilde{v}_2 \ \tilde{v}_3 \ \cdots \ \tilde{v}_k]^T$, applying Proposition 2 to $\tilde{\mathcal{L}}(\lambda)$ we obtain $\mathcal{L}(\lambda) = \mathcal{L}(\lambda)^B$. \square

Recalling that any block-symmetric matrix pencil $\mathcal{L}(\lambda) = \mathcal{L}(\lambda)^B$ from $\mathbb{M}_1(P)$ is in $\mathbb{DM}(P)$ we obtain

$$\mathbb{DM}(P) = \{ \mathcal{L}(\lambda) \in \mathbb{M}_1(P) \mid \mathcal{L}(\lambda) = \mathcal{L}(\lambda)^B \}.$$

Clearly, also all pencils in $\mathbb{DL}(P)$ for matrix polynomials $P(\lambda)$ in monomial basis are block-symmetric. This has first been proven in [6]. In [3] it is shown that if $P(\lambda)$ is a singular matrix polynomial of degree $k \geq 2$, then none of the pencils in $\mathbb{DL}(P)$ is a linearization of $P(\lambda)$. Different, larger vector spaces of block-symmetric strong linearizations of matrix polynomials in monomial basis have been proposed in [2].

5. Construction of block-symmetric Pencils

This section is dedicated to the construction of pencils in $\mathbb{DM}(P)$ for regular $P(\lambda)$. It turns out that the characterization (5) yields a simple procedure to construct block-symmetric pencils. Once more assume $P(\lambda)$ to be of the form (2) with $\deg(P(\lambda)) \geq 2$. Moreover, let $\mathcal{L}(\lambda)$ be an element of $\mathbb{M}_1(P)$ as in (9), i.e.

$$\mathcal{L}(\lambda) = [(v \otimes I_n) \ B] F_\Phi^P(\lambda) = [(v \otimes \alpha_{k-1}^{-1} P_k) \ B] \lambda + [(v \otimes I_n) \ B] F_\Phi^P(0).$$

Similar to the block-skew-symmetric case, $[(v \otimes \alpha_{k-1}^{-1} P_k) \ B]$ being block-symmetric implies

$$B = \begin{bmatrix} Z \\ B^\star \end{bmatrix}$$

with $Z = \alpha_{k-1}^{-1} P_k \otimes [v_2 \ v_3 \ \dots \ v_k]$ and a $(k-1)n \times (k-1)n$ block-symmetric matrix B^\star . Therefore, it suffices to compute the lower block-trigonal part⁴ of B^\star and to only consider $[(v \otimes I_n) \ B] F_\Phi^P(0) = \mathcal{L}(0)$ for the remaining derivations. The block-symmetry certainly requires

$$(e_i \otimes I_n) \mathcal{L}(0) (e_1 \otimes I_n) \stackrel{!}{=} (e_1 \otimes I_n) \mathcal{L}(0) (e_i \otimes I_n). \quad (10)$$

As the first block column $(v \otimes I_n)$ and the first block row Z of $\mathcal{L}(0)$ are already known, equation (10) reads for $2 \leq i \leq k$

$$\begin{aligned} v_i \left(-\frac{\beta_{k-1}}{\alpha_{k-1}} P_k + P_{k-1} \right) - \alpha_{k-2} B_{i,1} \\ \stackrel{!}{=} v_1 P_{k-i} - \frac{(v_{i-1} \gamma_{k-i+1} + v_i \beta_{k-i} + v_{i+1} \alpha_{k-i-1})}{\alpha_{k-1}} P_k \end{aligned}$$

whereby we set $v_{k+1} = \alpha_{-1} = 0$ for $i = k$.⁵ This can easily be solved for the matrix $B_{i,1}$ and yields

$$\begin{aligned} B_{i,1} = & \frac{(v_{i-1} \gamma_{k-i+1} + v_i (\beta_{k-i} - \beta_{k-1}) + v_{i+1} \alpha_{k-i-1}) P_k}{\alpha_{k-1} \alpha_{k-2}} \\ & + \frac{(v_i P_{k-1} - v_1 P_{k-i})}{\alpha_{k-2}}. \end{aligned} \quad (11)$$

In this way, the blocks $B_{i,1}$ can be computed for all $i = 2, \dots, k$. Due to the block-symmetry of B^\star this completely and uniquely determines the first block column and block row of B^\star . In the same way, considering

$$(e_i \otimes I_n) \mathcal{L}(0) (e_2 \otimes I_n) \stackrel{!}{=} (e_2 \otimes I_n) \mathcal{L}(0) (e_i \otimes I_n).$$

gives the equation

$$\begin{aligned} v_i \left(P_{k-2} - \frac{\gamma_{k-1}}{\alpha_{k-1}} P_k \right) - \beta_{k-2} B_{i,1} - \alpha_{k-3} B_{i,2} \\ \stackrel{!}{=} v_2 P_{k-i} - (\gamma_{k-i+1} B_{2,i-2} + \beta_{k-i} B_{2,i-1} + \alpha_{k-i-1} B_{2,i}). \end{aligned}$$

⁴Notice that, in terms of B , it holds that $B_{s,t} = B_{t+1,s-1}$ for $t < s$ and $s \geq 2$.

⁵Terms involving α_{-1} will show up for $i = k$ in subsequent formulas, too. We will always assume $\alpha_{-1} = 0$.

It follows from the block-symmetry of B^\star that $B_{2,i-2} = B_{i-1,1}$, $B_{2,i-1} = B_{i,1}$ and $B_{2,i} = B_{i+1,1}$. Thus we obtain an explicit expression for $B_{i,2}$:

$$B_{i,2} = \frac{(\gamma_{k-i+1}B_{i-1,1} + (\beta_{k-i} - \beta_{k-2})B_{i,1} + \alpha_{k-i-1}B_{i+1,1})}{\alpha_{k-3}} + \frac{(v_i P_{k-2} - v_2 P_{k-i})}{\alpha_{k-3}} - v_i \frac{\gamma_{k-1}}{\alpha_{k-3}\alpha_{k-1}} P_k. \quad (12)$$

Due to the block-symmetry it suffices to consider (12) only for $i \geq 3$. Therefore, the blocks $B_{3,2}, \dots, B_{k,2}$ may be computed via (12). Following the same pattern, the equation $(e_i \otimes I_n)\mathcal{L}(0)(e_j \otimes I_n) = (e_j \otimes I_n)\mathcal{L}(0)(e_i \otimes I_n)$ yields in its most general form for $j \geq 3$ and $i \geq j$

$$B_{i,j} = \frac{\gamma_{k-i+1}B_{i-1,j-1} + (\beta_{k-i} - \beta_{k-j})B_{i,j-1} + \alpha_{k-i-1}B_{i+1,j-1} - \gamma_{k-j+1}B_{i,j-2}}{\alpha_{k-j-1}} + \frac{(v_i P_{k-j} - v_j P_{k-i})}{\alpha_{k-j-1}}. \quad (13)$$

Hence, we may interpret the blockwise computation of B as some kind of updated recurrence relation. Moreover, the derivation shows that (11) - (13) are sufficient and necessary for $\mathcal{L}(\lambda) \in \mathbb{DM}(P)$ being block-symmetric and having ansatz vector v .

We summarize the procedure to compute block-symmetric pencils in $\mathbb{M}_1(P)$: For any regular matrix polynomial $P(\lambda) \in \mathbb{R}^{n \times n}$ expressed in some orthogonal basis as in (2) and of degree $k \geq 2$ choose any $v \in \mathbb{R}^k$ and compute

$$B = \begin{bmatrix} Z \\ B^\star \end{bmatrix} \in \mathbb{R}^{kn \times (k-1)n} \quad B = [B_{i,j}], B_{i,j} \in \mathbb{R}^{n \times n}$$

according to (11) - (13) and set $Z = \alpha_{k-1}^{-1} P_k \otimes [v_2 \ v_3 \ \dots \ v_k]$. Then

$$\mathcal{L}(\lambda) = [(v \otimes I_n) \ B] F_\Phi^P(\lambda)$$

is block-symmetric with ansatz vector v .

Notice that this algorithmic approach does not require the computation of a single matrix-matrix-multiplication, only scalar-matrix-multiplications are needed. The complexity of this procedure is $\mathcal{O}(k^2 n^2)$, which also is the complexity of the construction algorithm presented in [9, Sec. 7]. Although there are structural similarities between both algorithms, they rise from quite different viewpoints.

Fortunately, now we obtain the following corollary without real effort.

Corollary 4. *For any regular $P(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ of degree k , $\dim(\mathbb{DM}(P)) = k$.*

Corollary 4 follows from the fact, that $\dim(\mathbb{DM}(P)) \geq k$ certainly holds because $\mathcal{B}_1(\lambda), \dots, \mathcal{B}_k(\lambda) \in \mathbb{DM}(P)$ with $\mathcal{B}_j(\lambda) = [(e_j \otimes I_n) \ B_j] F_{\Phi}^P(\lambda)$ are obviously linear independent. Moreover, since $\mathcal{L}(\lambda) = \sum_{i=1}^k \alpha_i \mathcal{B}_i(\lambda)$ for any $\alpha_i \in \mathbb{R}$ is block-symmetric with ansatz vector $v = \sum_{i=1}^k \alpha_i e_i$, whenever any $\mathcal{L}^*(\lambda) \in \mathbb{DM}(P)$ has the same ansatz vector, we necessarily have $\mathcal{L}(\lambda) = \mathcal{L}^*(\lambda)$ due to the uniqueness of the expressions (11) - (13). Thus $\dim(\mathbb{DM}(P)) \leq k$ and Corollary 4 follows.

6. Conclusion

We presented a rigorous generalization of the results obtained in [8] to orthogonal polynomial bases. Although the extension of the concepts from [8] to nonstandard bases has already been considered, it was one of our main aims to present the subject in a coherent and concise manner introducing some new aspects without drawing on deeper theoretical results. Setting up the generalized ansatz spaces as introduced in [8, Sec. 4.2], we were able to characterize the elements in these spaces nicely, obtain a simple linearization condition and prove statements on the space dimension or the genericity of linearizations without any effort. Moreover, we presented a basic and short algebraic proof on the fact that double generalized ansatz spaces contain entirely block-symmetric pencils using a rather surprising argument on block-skew-symmetric pencils. In the last section, we derived an intuitive procedure to construct block-symmetric pencils in generalized ansatz spaces.

Acknowledgements

The second author would like to thank D. Steven Mackey and Froilán M. Dopico for their friendly remarks on this work at the 2016 ILAS conference in Leuven, Belgium.

References

- [1] Amiraslani, A., Corless, R., Lancaster, P., 2009. Linearization of matrix polynomials expressed in polynomial bases. IMA J. Numer. Anal. 29 (1), 141–157.

- [2] Bueno, M., Dopico, F., Furtado, S., Rychnovsky, M., 2015. Large vector spaces of block-symmetric strong linearizations of matrix polynomials. *Linear Algebra Appl.* 477, 165–210.
- [3] De Terán, F., Dopico, F., Mackey, D., 2009. Linearizations of singular matrix polynomials and the recovery of minimal indices. *Electron. J. Linear Algebra* 18, 371–402.
- [4] Effenberger, C., Kressner, D., 2012. Chebyshev interpolation for nonlinear eigenvalue problems. *BIT Numer. Math.* 52 (4), 933–951.
- [5] Faßbender, H., Saltenberger, P., 2016. Some notes on the linearization of matrix polynomials in Tschebyscheff basis. doi:10.13140/rg.2.1.1044.4888.
- [6] Higham, N. J., Mackey, D. S., Mackey, N., Tisseur, F., 2006. Symmetric linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.* 29 (1), 143–159.
- [7] Lawrence, P., Perez, J., 2016. Constructing strong linearizations of matrix polynomials expressed in the Chebyshev basis. MIMS EPrint 2016.12, Manchester Institute for Mathematical Sciences, The University of Manchester, UK.
- [8] Mackey, D., Mackey, N., Mehl, C., Mehrmann, V., 2006. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.* 28 (4), 971–1004.
- [9] Nakatsukasa, Y., Noferini, V., Townsend, A., 2012. Vector spaces of linearizations for matrix polynomials: a bivariate polynomial approach. MIMS EPrint 2012.118, Manchester Institute for Mathematical Sciences, School of Mathematics, The University of Manchester, Manchester, Great Britain.
- [10] Noferini, V., Pérez, J., 2015. Chebyshev-Fiedler pencils. MIMS EPrint 2015.90, Manchester Institute for Mathematical Sciences, The University of Manchester, UK.
- [11] Robol, L., Vandebril, R., Dooren, P. V., 2016. A framework for structured linearizations of matrix polynomials in various bases. arXiv:1603.05773v2.